Effect of Title Content on Clicking Behavior

Hulse, J.D., Specht, A.R., Bowling, A., & Buck, E.

The Ohio State University

Introduction

In the modern age, the internet has become one of the most common resources for information gathering (Nie & Erbring, 2002). The internet creates a unique environment "where information sources and channels proliferate, and people have a high level of control over what to choose" causing the influence of selective exposure to be emphasized (Kim, Forquer, Rusko, Hornik & Cappella, 2016, p. 1). Much research has been done on the topic of selective exposure and how it relates to how people choose which articles to read and, by extension, what information to consume. The general concept underlying selective exposure is that people will select what information to take in based on certain preexisting factors or biases. One of the main factors is the avoidance of psychological stress, such that people choose to expose themselves to information that aligns with their existing beliefs or biases in order to avoid experiencing cognitive dissonance, which is an uncomfortable state of psychological stress caused by holding two or more contradicting beliefs. Originally, selective exposure theory mainly predicted the avoidance of information that could cause cognitive dissonance (Knobloch, Carpentier & Zillmann, 2003, Sears & Freedman, 1967). However, over the years, the concept of selective exposure has expanded from the study of the avoidance of cognitive dissonance to encompass the study of other factors driving the active selection of certain information (Knobloch et al, 2003).

There are many seemingly small factors that can influence what articles people actively choose to read and what information they actively expose themselves to. Factors that can influence such active selective exposure include a variety of matters from utility of the information to social considerations. The concept of informational utility suggests that "messages conveying more useful or practical information are more likely to be selected than less useful or practical messages" (Kim et al, 2016, p. 2). Social considerations can reference the sum of a person's relationships within society, according to the co-orientational model, which operates on the idea that "person's use of mass communications does not occur in isolation from the rest of his social life" but rather is "interwoven in an outgoing system of reciprocal influence" (Chaffee & Mcleod, 1973, p. 237). Social considerations can also reference a variety of demographic influences affecting information selection, such as race, gender or political background. Other factors include things such as negativity bias, meaning people are more likely to read news that is framed in a negative way, or position bias, meaning people's choice of article is influenced by where it is physically presented on the site (Kim et al, 2016).

One such factor that has been examined in multiple studies is the effect of source quality on selective exposure. One would generally assume that given the choice between information from a high-quality source and a low-quality source, people would choose to expose themselves to the information from the high-quality source. However, counterintuitively, it has been found that source, where the source is considered to be the site on which the information originates, has very little influence on selective exposure. Studies where quality of source was

indicated in the source name found that while people could distinguish between a good and bad source when source names were presented in isolation, when it came time to select articles, the source did not affect which articles were selected (Westerwick, Johnson & Knobloch-Westerwick, 2016).

The effect of source quality as indicated through headlines has also been looked into. For example, a study done on selective exposure to health information, in which the quality of the source was indicated in the title of the article, found that "headlines that explicitly indicated expert information sources did not differ in the number of audience selections from those without such sources" (Kim et al, 2016, p. 13). An interesting caveat is that source has been found to have an influence when regarded in a social context – "influential users" on social media do have influence as a 'source' (Kwon, Han & Kim, 2017) even though they are simply sharing/promoting information. The source of the information itself is still not influential, but the social standing of the person sharing the information on social media does have an impact (Kwon et al, 2017). A numeric indication of popularity of a specific article on a news site has also shown to influence selection (Knobloch-Westerwick, Sharma, Hansen & Alter, 2005).

Content of headlines can also influence which articles an audience selects. When people are browsing magazines or websites, whether or not a headline catches their attention is often how people decide what articles to read. The content of headlines can even be manipulated to produce specific emotions that cause people to be more drawn to one article over another (Swinehart, 1968). In recent years, there has been a rise in the occurrence of sensationalist and clickbait headlines and that are designed to get people to read an article rather than to accurately convey information (Frampton, 2015), so it is becoming increasingly important to understand how and why these articles are being selected.

As previously discussed, a wide range of factors affect what articles people select to expose themselves to. Most studies tend to observe the effect of specific aspects of the way the message is being presented or the participants mindset and political views, so there are some other factors that have not been the subject of much study. Notably, one potential factor of which there seems to be little to no investigation is the reader's level of education and how that might affect selection. Whether there is any potential correlation between education level and the quality of information which is selected has not yet been determined. Other demographic factors such as age, race, and gender are often recorded in studies on selective exposure as the effects of different heuristics can vary among different groups of people because of the differences in values, life experiences, world views, et cetera that exist between demographics. It could be beneficial to look into how level of education interacts with selective exposure, especially in the context of selecting based on headlines and source quality, as those factors may be more influential on people who have presumably learned to look for them through the course of their higher education compared to people who have not learned to discern between information sources.

This study focuses on selective exposure in the realm of health and human diet. The problem this study focuses on is what factors affect which online article concerning diet and/or health information people decide to read. Many studies have shown a correlation between higher level of education and better health (Winkleby, Jatulis, Frank & Fortmann, 1992), so this study will serve to see if higher education influences health via what information is consumed. More specifically, this study intends to determine how the content of headlines influences clicking behavior. As a secondary factor, this study looks into how level of education affects which articles are selected. Very little research has been done on how level of education factors into selective exposure, so this study observes whether there is any correlation between the education level of participants and the quality of health-related information to which they expose themselves. Presumably, more educated people will have learned how to select credible sources of information because classes often require credible sources to be cited in papers, so it would be interesting to see if learning how to make good selections actually leads to making good selections.

This study has three objectives:

(Obj.1) To determine if content source will affect selection;

(Obj.2) To determine if articles with clickbait style titles, defined by their emphasis on grabbing a reader's attention or sensationalism rather than factual reporting, will have a higher rate of selection overall;

(Obj.3) To determine if students with higher level of education will select for more reliable sources based on title more often than students with lower level of education.

Procedures and Methods

Population and Sampling

The population for the study was young adults enrolled in a four-year university, specifically The Ohio State University College of Food, Agricultural, and Environmental Sciences since they were able to be contacted in large numbers through the college's newsletters and listservs. The study's sample was determined by the subjects self-selecting to participate in the study. Participants were offered the chance to enroll in a gift card drawing as an incentive to participate. The survey was sent to a total of 2,890 students. From that population, a sample of 128 responses were received. There were 2 responses from high school seniors (1.7% of responses), 35 from college freshmen (30.4% of responses), 16 from sophomores (13.9% of responses), 21 from juniors (18.3% of responses), 34 from seniors (29.6% of responses), and 7 from people who had graduated college (6.1% of responses). 78.1% of respondents were female, 11.7% of respondents were male, and 10.2% of respondents did not provide gender demographic information. 78.9% were white, 5.4% of respondents were Asian/Pacific Islander, 3.1% of

respondents were Black/African American, 6.3% of respondents were Hispanic/Latino, and 6.3% of respondents did not provide race demographic information. The sample is certainly not representative of the population at large but is in some ways representative The Ohio State University. While the data skewed in terms of gender it was fairly accurate in terms of race. As of 2018, the percentage female students enrolled at Ohio States main campus was approximately 50.3% and the percentage of white students was approximately 79% (Statistical Summary 2018-2019). Demographic information on grade level distribution and specific demographic information for the College of Food, Agricultural, and Environmental Sciences was not available for analysis.

Instrumentation

This study utilized an online survey, which took participants approximately five minutes to complete. When they began the survey, the participants were given shown a consent form and confirmed they were over eighteen years of age. They then moved on to a screen where they were asked to select an article from four choices (Fig 1). The four articles were related to health and diet and had differing qualities of title and source: one with a formal title and a reliable source, one with a formal title and an unreliable source, one with a clickbait style title and a reliable source, and one with a clickbait style title and an unreliable source. For the purposes of this study, clickbait was considered to be titles that were geared toward grabbing attention with short statements like “diet or dud?” or suggesting that the article is a listicle as with statements like “6 types … and what to know about them.” The titles considered more scientific for the purpose of this study were those that were more informative about the specific topic being discussed in the article.

The participants were told to simply click the article they would most like to read based on the provided title, description, and source information. Care was taken to ensure that all the stories discussed the same topic – dieting – and varied only in the way the topic was presented. To ensure the instrument had face validity, all article titles and description were real articles taken from real news sources. Participants were asked demographic information after completing the question in order to try to limit the salience of their demographics in their article choices.

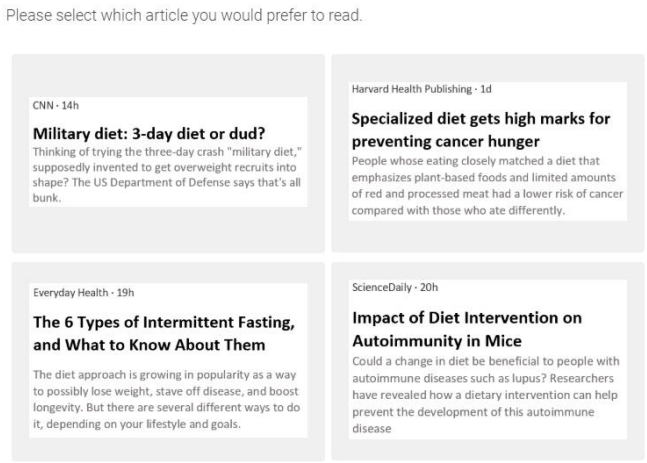

Figure 1: The stimulus as presented in the study: clickbait style title/reliable source (top left), formal title/reliable source (top right), clickbait style title/unreliable source (bottom left), formal title/unreliable source (bottom right)

Data Analysis

We measured the frequency of clicking behavior on different type of health-related stories and the potential impact of education level on that clicking behavior. We used quantitative measures such as frequencies and a Pearson Chi Squared test for independence to determine if level of education influenced the likelihood of participants clicking on different types of stories. We also observed what stories were selected by different grade levels to see if there was a relationship between grade level and story. Data analysis was done using Qualtrics Stats iQ function.

Results

Out of the 128 survey responses received, 117 people selected articles. This provided a completion rate of 91.4%. The results supported Objective 1 but did not support Objective 2 or Objective 3.

The results do provide support for (Obj.1); the source of the content did not appear to affect selection. Both the two most selected articles and the two least selected articles had one reliable source and one unreliable source among them.

Instead of finding that (Obj.2) articles with clickbait style titles had a higher rate of selection, results showed no relationship between title style and selection. The two articles selected most frequently were the one with a clickbait title and an unreliable source and the one with a scientific title and a reliable source (Fig. 2). The clickbait title and unreliable source had 39 selections (33.33% of all selections). The scientific title and reliable source had 36 selections (30.77% of all selections). The clickbait title with a reliable source had 18 selections (15.38% of all selections) and the scientific title with the unreliable source had 24 selections (20.51% of all selections).

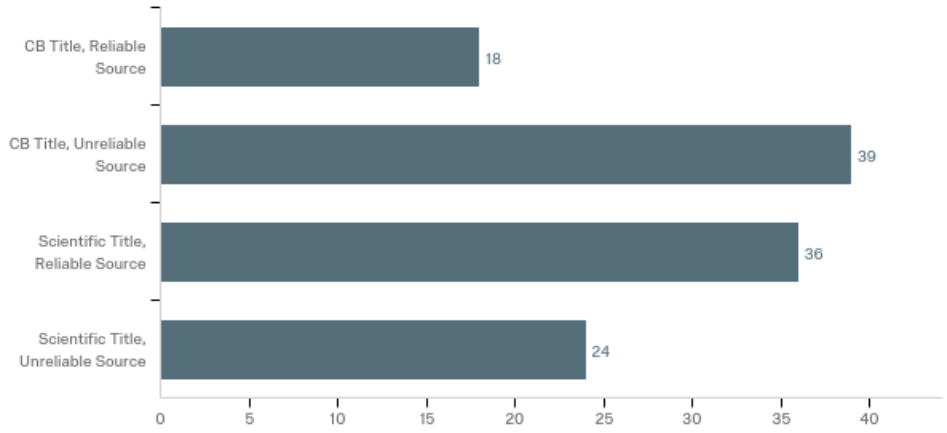

Figure 2: showing the number of times each article was selected by participants

Instead of (Obj.3) students with higher level of education selecting more scientific title more often than students with lower level of education, results showed no relationship between grade level and selection or between years of college completed and education. The factor of education level was examined by comparing the selected articles of university freshman, who only have completed a high school education, to seniors, who have completed most of their college education. The Pearson Chi Squared test for independence showed no significant relationship between grade level and article selected (Fig. 3) ($\chi^2 = 0.190, p = 0.644$). Similarly, the Pearson Chi Squared test for independence showed no significant relationship between years of college completed and article selected (Fig. 4) ($\chi^2 = 0.164, p = 0.861$). Both grade level and number of years of college completed were tested for separately to account for the potential difference between incoming freshmen, who would have completed 0 years of college, and freshmen at the end of their first year, who would have completed 1 year of college.

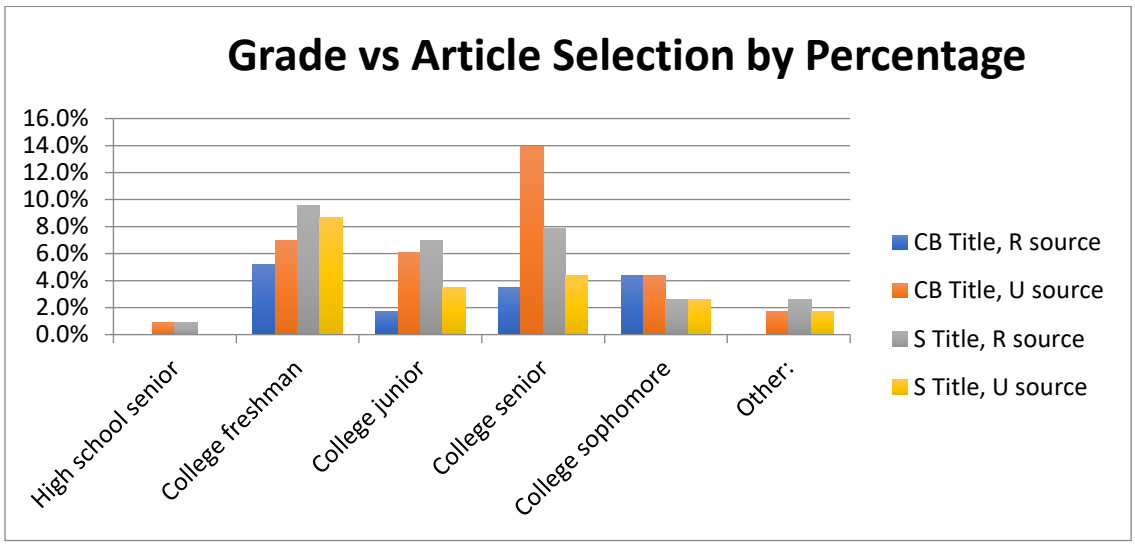

Figure 3: Grade vs Article Selection by Percentage

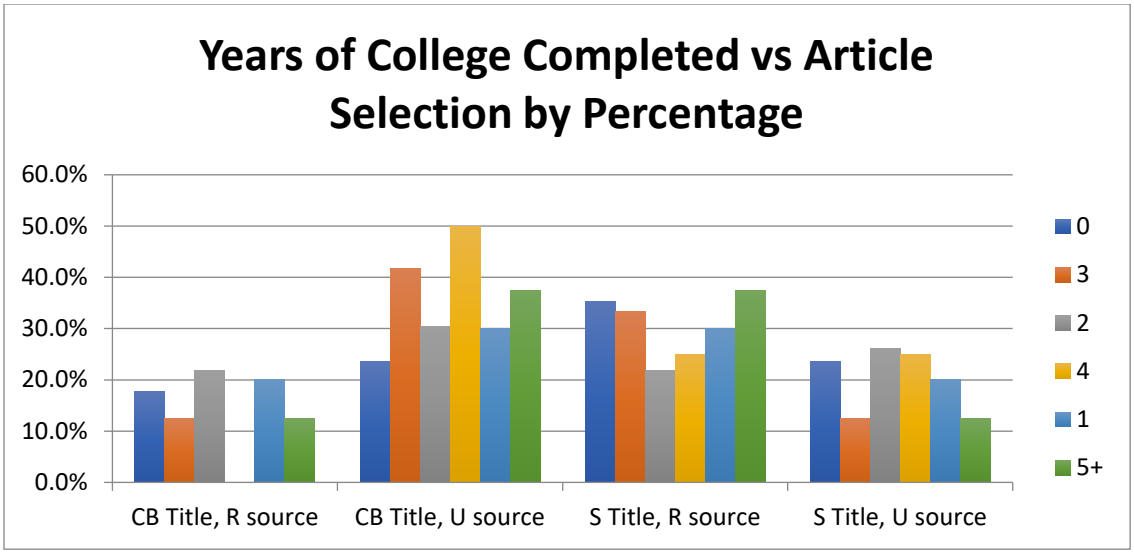

Figure 4: Years of College Completed vs Article Selection by Percentage

Discussion

While the results did confirm that the source did not appear to influence the selection of articles, interestingly the wording of the title did not appear to affect the results of this study either. This may have been because there were only four articles to choose from. This was originally done to prevent any errors of habituation where participants simply choose at random in order to complete the study, but a replication of this study might consider having more articles to choose from in each category (the combinations of reliable/unreliable source and clickbait/scientific title), so that participants selecting for a specific topic that interests them does not influence the results. This study did not have time to preform pretesting of the survey instrument, so future studies would benefit from doing that to ensure reliability and validity.

Furthermore, the specific topics of the articles could have acted as a confounding variable. The two articles most selected for have topics that could be considered generally interesting to the public, whereas the specific topics of the two least selected articles might be aimed at more specific audiences. The least selected articles were about a military diet and a diet affecting autoimmunity in mice, respectively, and it would be understandable if these specific topics were simply less appealing to most people compared to the more heavily selected articles whose topics could be considered applicable to a wider audience. Interestingly, these results do appear to be in keeping with the idea of informational utility discussed in the introduction, although that is not what this study originally aimed to test for. This study initially aimed for a degree of authenticity by selecting article titles from real sources, but it would likely be easier for any future replications to control this confounding variable by writing fake article titles specifically for the study which are more closely matched in potential interest level.

Another shortcoming of the current study is the relatively small and homogenous pool of participants. It is difficult to pick out statistically significant trends among such a small number of responses received; a replication of this study should aim to have a bigger and more diverse sample size. It would be beneficial to have a sample with a wider range of education level than just college freshmen to college seniors. The closeness of these two groups may have obscured the effect education might have on selection. Although this study was not successful in finding any correlation between education level and selective exposure to health information, it could still be a beneficial route for future studies to pursue since it has largely gone untested thus far.

References

Chaffee, S. H., & Mcleod, J. M. (1973). Individual vs. Social Predictors of Information Seeking. *Journalism Quarterly, 50*(2), 237-245. doi:10.1177/107769907305000204

Frampton, B. (2015, September 14). Clickbait: The changing face of online journalism. Retrieved from https://www.bbc.com/news/uk-wales-34213693

Kim, H. S., Forquer, H., Rusko, J., Hornik, R. C., & Cappella, J. N. (2016). Selective Exposure to Health Information: The Role of Headline Features in the Choice of Health Newsletter Articles. *Media Psychology, 19*(4), 614-637. doi:10.1080/15213269.2015.1090907

Knobloch, S., Carpentier, F. D., & Zillmann, D. (2003). Effects of Salience Dimensions of Informational Utility on Selective Exposure to Online News. *Journalism & Mass Communication Quarterly, 80*(1), 91-108. doi:10.1177/107769900308000107

Knobloch-Westerwick, S., Sharma, N., Hansen, D. L., & Alter, S. (2005). Impact of Popularity Indications on Readers Selective Exposure to Online News. *Journal of Broadcasting & Electronic Media, 49*(3), 296-313. doi:10.1207/s15506878jobem4903_3

Kwon, J., Han, I., & Kim, B. (2017). Effects of source influence and peer referrals on information diffusion in Twitter. *Industrial Management & Data Systems, 117*(5), 896-909. doi:10.1108/imds-07-2016-0290

Nie, N. H., & Erbring, L. (2002). Internet and Society: A Preliminary Report. *IT&Society, 1*(1), 275-283. Retrieved from http://www.itandsociety.org/

Sears, D. O., & Freedman, J. L. (1967). Selective Exposure to Information: A Critical Review. *Public Opinion Quarterly, 31*(2), 194. doi:10.1086/267513

Statistical Summary 2018-2019. (n.d.). Retrieved from https://www.osu.edu/osutoday/StatisticalSummary2018.pdf.

Swinehart, J. W. (1968). Voluntary exposure to health communications. *American Journal of Public Health and the Nations Health, 58*(7), 1265-1275. doi:10.2105/ajph.58.7.1265

Westerwick, A., Johnson, B. K., & Knobloch-Westerwick, S. (2016). Change Your Ways: Fostering Health Attitudes Toward Change Through Selective Exposure to Online Health Messages. *Health Communication, 32*(5), 639-649. doi:10.1080/10410236.2016.1160319

Winkleby, M. A., Jatulis, D. E., Frank, E., & Fortmann, S. P. (1992). Socioeconomic Status and Health: How Education, Income, and Occupation Contribute to Risk Factors for Cardiovascular Disease. *American Journal of Public Health, 82*(6), 816–820.